

Search for rare SM processes in the $\cancel{E}_T + b$ -jets signature at CDF

K. Potamianos

Purdue University, West Lafayette, IN, USA

The missing transverse energy (\cancel{E}_T) plus b-jets signature is very promising for searches for the Higgs boson or new physics. Indeed, \cancel{E}_T naturally arises from unidentified particles such as neutrinos, neutralinos, gravitons, etc., and b-quarks are the main decay products of a low mass Higgs boson as well as of several exotic particles. The main challenge is to identify and reject the numerous standard model (SM) backgrounds that mimic this signature. This is especially so for QCD multi-jet production, a large background due to mis-measurement (rather than undetectable particles). We present state-of-the-art data-driven and multivariate techniques to characterize and reject this instrumental background. These techniques make analyses in this signature as sensitive as those using lepton identification and allow probing for rare SM processes. We describe searches for electroweak single top production, a part of the observation of single top by CDF, and for a low mass SM Higgs boson, one of the most sensitive among low mass Higgs searches at CDF. We also present a measurement of the top pair cross-section in this signature, and discuss other analyses and future prospects.

1 Introduction

The standard model of particle physics (SM) accurately describes most physical observables not involving gravity. All of its particles have been observed and characterized, save to the Higgs boson¹, whose role is to provide mass to the elementary particles. Direct searches and precise electroweak fits constrain its mass, and favor a low mass ($m_H < 158 \text{ GeV}/c^2$ at 95 % C.L.)².

At the Tevatron, the low mass range is best investigated through the associated production of a Higgs boson and a W or Z boson, whose leptonic decay products are triggered on³. At low mass, the Higgs predominantly decays to a $b\bar{b}$ pair. In the following, we focus on $ZH \rightarrow \nu\nu b\bar{b}$ events. These yield a large transverse energy imbalance (\cancel{E}_T), two high- p_T b -jets, and no identified lepton. Due to the difficulty to identify and reconstruct leptons, we are also sensitive to $WH \rightarrow \ell\bar{\ell} b\bar{b}$. This is true especially when the W boson decays to a τ that is not reconstructed ($\sim 50\%$ of the times). Finally, we accept a tiny fraction of $ZH \rightarrow \ell\bar{\ell} b\bar{b}$ events.

In addition to Higgs physics, this signature is also sensitive to the electroweak production of a single top quark, diboson production, and top pair production. These SM processes are either categorized as signal or background, while events from QCD multi-jet production and from the production of a vector boson in association with jets are part of the background.

Next to events from SM processes, many models of new physics also predict this signature: SUSY ($b\bar{b} \rightarrow b\tilde{\chi}^0\bar{b}\tilde{\chi}^0$), technicolor ($\rho_T^\pm \rightarrow Z\pi_T^\pm \rightarrow \nu\nu b\bar{q}$ and $\rho_T^\pm \rightarrow W^\pm\pi_T^0 \rightarrow \ell\nu b\bar{b}$), extra-dimension, etc. In this proceeding, we focus only on SM physics.

In the $\cancel{E}_T + b$ -jets signature, the signal is very small when compared to the large backgrounds, especially from QCD production. In the following, we present an innovative technique to isolate and reject the large QCD background.

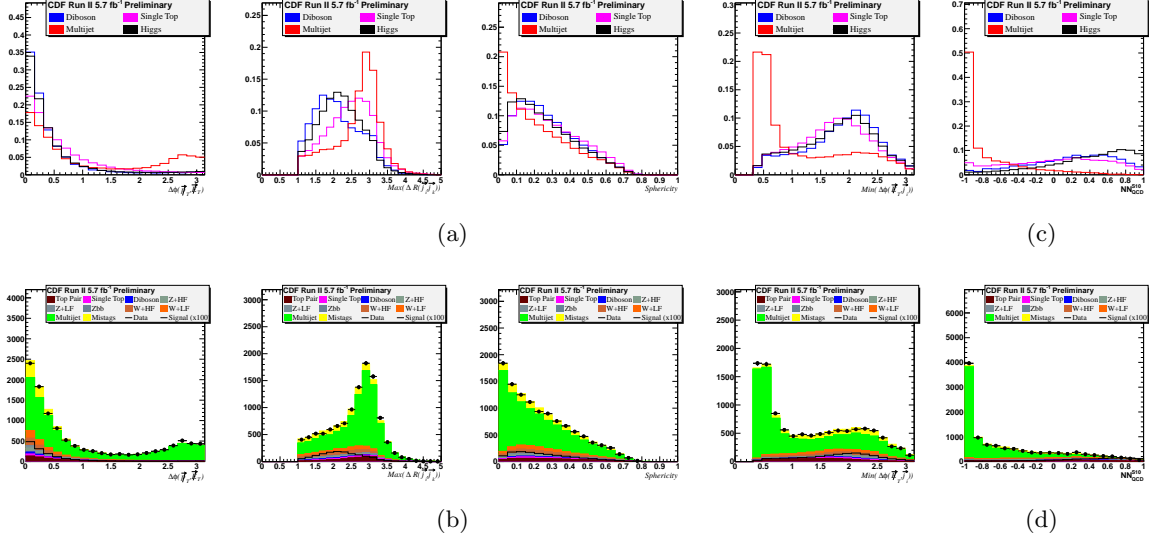


Figure 1: Some of the input variables to the neural network used to reject the QCD multi-jet background: (a) shape comparison, and (b) model validation. (c), (d): Neural network output.

2 Background modeling

The $\cancel{E}_T + b$ -jets signature accepts events from every SM process. We use a combination Monte Carlo and data-driven techniques to provide a proper model for each of them. The electroweak processes yield *real* \cancel{E}_T coming from the neutrino(s) in the event. We use Monte Carlo to model these⁴. Additionally, we use data to predict the rate at which light flavor events from these processes can be wrongly identified as originating from a b -quark (i.e. b -tagged). On the contrary, QCD multi-jet production (MJ) yields *instrumental* \cancel{E}_T , due to the mis-measurement of the energy of the jets. Because of a large cross-section and of the presence of non-negligible high order effects, it is impractical and not advised to use Monte Carlo to model MJ, especially with large datasets. We therefore use a data-driven model that defines a four-dimensional matrix to predict the probability for a data event to be b -tagged: the *Tag-Rate-Matrix*⁴.

3 Tools for background rejection

We select events with $\cancel{E}_T > 50 \text{ GeV}/c^2$, two or three jets ($> 15 \text{ GeV}/c^2$, $|\eta| < 2$), and no identified lepton. We further require the leading (second) jet to have an energy of at least 35 (25) GeV/c^2 and either one to be central ($|\eta| < 0.9$). We remove mis-measured QCD events requiring $\Delta\phi(\cancel{E}_T, j_{(2,3)}) > 0.4$ and $\Delta\phi(\cancel{E}_T, j_1) > 1.5$. We then further improve the S/B ratio.

3.1 Signature of missing particles

A common technique to identify a particle such a neutrino is to measure the transverse energy imbalance in the calorimeter (\cancel{E}_T). Here, we also rely on an *independent* sub-detector, the spectrometer, to determine the transverse momentum flow imbalance (\cancel{p}_T). In the case of a particle escaping the detector, the \cancel{E}_T and the \cancel{p}_T are aligned. However, for a mis-measured QCD di-jet event, the \cancel{E}_T and is either aligned or anti-aligned to the \cancel{p}_T (Figure 1a, left).

3.2 A neural network to reject the QCD multi-jet background

Next to comparing \cancel{p}_T to \cancel{E}_T , we identify kinematic quantities discriminating the signal from the QCD background. Instead of cutting on each variable, we feed a neural network that exploits their correlations, and then cut on its output, reducing the background more efficiently.

We train our network using the *pre-tag* data sample weighted by the *Tag-Rate-Matrix* to model MJ events. This choice is preferred over the use of a Monte Carlo QCD background because it allows not only to reject QCD but also a part of the other electroweak backgrounds. With this technique, it is possible to reject about 90% (70%) of the MJ (overall) background while only losing 10% of the signal. The signal sample is composed of Monte Carlo events in proportion to the relative size of each signal component.

Figure 1a shows the difference in shape for some of the network input variables. Each of the variables is validated against data, as shown in Figure 1b. In addition to the region where the network is derived, we use five other control regions in which we check both the inputs to the network and its output. This makes us confident in using multivariate analysis to find a signal.

4 Analyses using this neural network background rejection technique

The technique we present is very generic and can be trained with different signals. We briefly describe several analyses that owe their significance to this technique. After rejecting the QCD background, we train another neural network to discriminate the signal from the background.

4.1 Measurement of the top pair production cross-section

First observed in 1995⁵, the top quark has been extensively studied at the Tevatron, mostly in the (semi-)leptonic and all-hadronic signatures. In 2010, we measured for the first time the top pair production cross-section in the $\cancel{E}_T + b$ -jets signature⁶. This measurement is complementary to those preceding and contributes to improving the world average. Much importantly, probing a well known signal is a stringent test of the analysis technique. Analyzing 5.7fb^{-1} of CDF data, we measure a cross-section of $7.12^{+1.20}_{-1.12}$ pb, assuming $m_t = 172.5 \text{ GeV}/c^2$. This measurement is as sensitive as the dilepton and all-hadronic signatures (Figure 2a).

4.2 Measurement of the single top production cross-section

In 2009, CDF and DZero observed the electroweak production of a single top quark⁷. The single top production cross-section is directly proportional to the square of the $|V_{tb}|$ element of the CKM matrix⁸ and a measurement thus constrains fourth-generation models. Our analysis in the $\cancel{E}_T + b$ -jets signature⁴ contributed to the result by adding 30% orthogonal signal into the combination^{7,9}. Analyzing 2.1fb^{-1} of CDF data, we measure a cross-section of $4.9^{+2.5}_{-2.2}$ pb (observed sensitivity of 2.1σ) and a $|V_{tb}|$ value of $1.24^{+0.34}_{-0.29} \pm 0.07$ (theory), assuming $m_t = 175 \text{ GeV}/c^2$ (Figure 2b).

4.3 Search for the SM Higgs boson

The $\cancel{E}_T + b$ -jets signature is one of the most sensitive for probing a low mass SM Higgs boson, and an important component of the CDF and Tevatron combination³. This search is challenging due to the tiny signal. We tune the analysis to $ZH \rightarrow \nu\nu b\bar{b}$ and $WH \rightarrow \ell\nu b\bar{b}$ ^{10,11}. Analyzing 5.7fb^{-1} of CDF data, we set a limit below 5 times the standard model prediction for a Higgs mass up to $135 \text{ GeV}/c^2$, and at 3 times for $m_H = 115 \text{ GeV}/c^2$ (Figure 2b).

5 Summary and future prospects

The background rejection technique presented here yields similar performance to lepton identification and allows to exploit the large acceptance of the $\cancel{E}_T + b$ -jets signature. We have presented several analyses using this innovative technique to isolate small signals. Further steps include relaxing the pre-selection requirements (lower \cancel{E}_T and jet energy cuts) to gain in acceptance while still rejecting the large backgrounds. We then plan to measure the diboson production cross-section in this signature.

